Plan for Analyzing and Testing Hypotheses (PATH)
Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook

Reviews by the
Scientific Review Panel

Dr. Jeremy Collie
Dr. Jim Kitchell
Dr. Saul Saila
Dr. Carl Walters

Dr. Jeremy Collie

Reviewer: Jeremy S. Collie
Title of Paper: Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook
Author: Marmorek and Peters [eds.]

Comments: Overall this draft decision analysis report is well organized and well written. This report summarizes a large number of prospective analyses in a format that should be understandable to decision makers.

a) scientific soundness of the methodology

This report provides a useful framework for constructing a decision table for analyzing management decisions, primarily concerning the hydrosystem. An extended family of alternative hypotheses is analyzed across three management actions. At this stage, all hypotheses are weighted equally, so no comments are provided on which of the hypotheses are more likely. Instead, the authors conduct sensitivity analyses to see which of the uncertainties (sets of hypotheses) have the greatest influence on meeting the survival and recovery standards. "Once the weights are assigned, focus can then shift to defining those combinations of hypotheses that are internally logical and are consistent with specific retrospective hypotheses." (p. 80)

I fear that these steps are being done in the wrong order. I think that the first step should be to test whether hypotheses are internally logical and that they are consistent with the retrospective aggregate hypotheses. Hypotheses that fail this test should be assigned a low weight. There is a risk of circular reasoning if the sensitivity of outcomes to hypothesis weighting is known prior to assigning the weights. A person with a preferred management decision would tend to assign a higher likelihood to a hypothesis that implies the preferred action. Likelihoods should be assigned to hypotheses objectively and based on empirical data wherever possible. The hypothesis (state of nature) should not be confused with the management action which would be preferred if that hypothesis were true.

The methodology is pseudo-scientific. Scientific "factoids" are used selectively to support established beliefs. There is a growing credibility gap between the models and the observations upon which they are based. The PATH analysts are basically trying to construct Level III hypotheses from Level I data.

b) general suitability of the data for use in the analyses

I was surprised by the total reliance on the CRiSP and FLUSH passage models. I guess that some sort of passage model is needed to accommodate the various hypotheses related to downstream passage survival. However, apart from the comparison with empirical reach survival and PIT-tag detection estimates, neither passage model has been fully calibrated to empirical data. FLUSH seems to be based to a greater extent on empirical relationships between survival and transit time. The general approach is to assume that the passage models accurately estimate survival and to then partition the residual mortality in the spawner-recruit model to other causes. However, if the passage models are inaccurate, the residual mortality would also be miss specified and incorrectly attributed to other causes.

In formulating hypotheses and assigning them weights, the analysts should heed William of Occam and the principle of parsimony. A simpler hypothesis that is consistent with retrospective data should be weighted higher than a less parsimonious hypothesis. A working null hypothesis is that the various sources of mortality are independent and additive. If such a model can explain the data, it should be favored over a model with second-order effects and interactions.

Some of the hypotheses here do not seem to be supported retrospectively, and others were not addressed in the retrospective analyses. In particular, there is no mention in Appendix A of the alpha model having been fit to the spawner-recruit data. It is not useful to have two separate model formulations (alpha and delta) to describe the same hypotheses. It should be possible to describe viable hypotheses with a single model form. The delta formulation is preferable because it is linked more closely to the MLE parameter estimates.

I don't think you should bother with cyclic climate models. No statistical evidence is presented that 18.5 or 60 year cycles in the environmental variables stand out from the spectrum of variability in the northeast Pacific. Nor are there statistical relationships between the environmental variables and salmon survival. With cycles, it is necessary to specify both the period and phase, neither of which is known with confidence. Climatic conditions were assumed to have been good in the early 1990s (p. 39), but was chinook salmon survival correspondingly high in those years? As the decision analysis results were insensitive to future climate effects, it is not worth trying to model them in detail. Most environmental variables exhibit variability on a continuous range of timescales, which resembles a red spectrum (Collie & Spencer 1994). Such variability is inherently unpredictable but its general pattern (including regime shifts) can be simulated with a first-order random variable, the variance of which is estimated from climatic effects from 1950 to 1995. A related approach is to model climatic variability with a sum of sine waves (Steele & Henderson 1984, Spencer & Collie 1997).

To be tenable, the hypotheses on extra (Post-BONN) mortality need to explain the higher mortality of Snake River stocks compared with the down-river stocks. The hydro-related hypothesis can explain this difference. The stock-viability hypothesis should be testable by observing whether the proposed mechanisms (e.g. BKD), occur to a greater extent for Snake River stocks than downstream ones. The USGS study cited in Appendix A - page 105 in support of the BKD hypothesis does not seem to relate to BKD, and survival was lower for stressed fish. It is inappropriate to list depensatory predation as a possible source of irreversible mortality here. There is already a depensatory term (p) in the stock-recruitment model to account for this possibility. Depending on the strength of depensation, a stock can still recover from low abundance, but its rate of increase would be lower than without depensation.

The near synchronous responses of salmon populations throughout the North Pacific to equally broad-scale regime shifts, make it very unlikely that Snake River stocks would respond with a markedly different pattern than other Columbia River chinook stocks. The data shown in Figure A.3.3.3-1 do not provide a good test for regime shifts in the productivity of wild Snake River chinook salmon. Many factors are confounded in these data: wild and hatchery run sizes, harvest, hydro development and climatic effects. Therefore the assumption that run sizes would be similar during 1936-1947 and 1977-1992 may be invalid.

c) validity of inference and conclusions reached

It was surprising that the drawdown scenario was not able to meet the 24-year survival standard for a higher proportion of runs, given that all the models predict higher in-river survival for this action (Figure 5.8-1). One reason must be the delay in achieving these higher survival rates. If the various drawdown hypotheses are weighted equally, the expected time until juvenile survival equilibrates would be approximately 12 years. While this delay is not unrealistic, it explains why A3 does not meet the 0.7 survival standard more often. I think that the 48-year recovery standard and 100-year survival standard should also be examined to distinguish between management actions that would never allow the stocks to recover from those that would rebuild the stocks with sufficient time.

d) suggestions for improvements and extensions to the analytical approaches used

In the decision table (5.2-1) the weights could be applied directly to the probabilities that the number of spawners exceed the survival escapement level. This would avoid the problem of the probability being just slightly greater or less than the criterion, without having to assign each outcome a graduated score (p. 61). The expected ability of an action to meet some criterion can be calculated directly without the columns of 0s and 1s in Table 5.2-1 (see table attached below).

Other performance measures could be based on the equilibrium properties of each model (aggregate set of hypotheses). For each model it should be possible to calculate the equilibrium or expected number of spawners. Then test whether this number is greater than 0 and greater than the survival standard. How long does it take the stock to reach this equilibrium under deterministic conditions? Stochastic simulations could be run in parallel to test whether the mean of the stochastic simulations differs markedly from deterministic runs. If not, it could save a lot of simulation. At any rate, equilibrium analyses would be useful for distinguishing extinction hypotheses from rebuilding hypotheses.

Fuzzy arithmetic is useful for uncertain variables where only an expected value and range are known (e.g. effects of habitat on the Ricker α value, juvenile survival rate after drawdown). Dr. Saila has suggested fuzzy arithmetic as a potentially useful alternative to Monte Carlo simulations. Fuzzy arithmetic allows the uncertainty of fuzzy numbers to propagate through life-cycle models, providing an aggregate measure of uncertainty.

e) opportunities for integration of the different component analyses into an adaptive management approach

The decision table is a page out of Walters' (1986) book on Adaptive Management of Renewable Resources. If drawdown were to occur, it should be done in a "staircase" experimental design to prevent confounding with climatic regime shifts.

f) relative priorities for future work on these analyses

A high priority should be assigned to estimating the likelihoods of aggregate retrospective models. Some of the hypotheses consist of responses to actions for which there is little or no prior experience (e.g. drawdown). In these cases conditional probabilities can be used to partition the overall likelihood into a component that can be estimated from empirical data, and components that must be estimated from expert opinion.

References

Collie, J.S. and P.D. Spencer. 1994. Modeling predator-prey dynamics in a fluctuating environment. Can. J. Fish. Aquat. Sci. 51:2665-2672.

Spencer, P.D. and J.S. Collie. 1997. Patterns of population variability in marine fish stocks. Fish. Oceanogr. 6:188-204.

Steele, J.H. and E.W. Henderson. 1984. Modeling long-term fluctuations in fish stocks. Science 224:985-987.

Table 5.2-1:	Probability that the number of spawners exceed the survival escapement level			
Aggregate Hypothesis	A1	A2	A3	Weight
1	0.35	0.65	0.45	0.05
2	0.35	0.65	0.47	0.08
3	0.84	0.85	0.78	0.09
4	0.78	0.85	0.78	0.08
5	0.33	0.63	0.46	0.01
6	0.32	0.65	0.47	0.02
7	0.84	0.84	0.78	0.07
8	0.74	0.85	0.78	0.08
9	0.82	0.68	0.63	0.05
10	0.83	0.71	0.63	0.07
11	0.51	0.53	0.67	0.05
12	0.51	0.53	0.67	0.06
13	0.66	0.60	0.75	0.01
14	0.67	0.61	0.76	0.00
15	0.53	0.52	0.66	0.07
16	0.53	0.52	0.66	0.07
17	0.76	0.60	0.75	0.06
18	0.68	0.60	0.74	0.04
19	0.53	0.50	0.67	0.03
20	0.53	0.51	0.67	0.01
Weighted probability of exceeding the survival standard	0.64	0.68	0.67	1

Dr. Jim Kitchell

Reviewer: James F. Kitchell
Title of Paper: Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook
Author: Marmorek and Peters [eds.]

General Overview

This review offers a general assessment of the PATH effort. Although I do know some about the background and events that led to this project, I have not been directly involved in the PATH process since I served as a reviewer of the modeling work conducted up through 1993. Much has changed since then.

I am encouraged by the fact that the modeling effort has focused on two, alternative approaches. Those models have fundamental differences that can help reveal the most appropriate ways to proceed with management actions. I am impressed by the extent and diversity of analyses conducted in the past few years. I am pleased to see that the models are being used to address questions focused on the management alternatives. Other reviewers have the background and familiarity required to provide an incisive evaluation of the modeling effort. My comments and suggestions are directed to expanding the conceptual and analytical approaches that may be pursued in future work.

Accomplishment of Objectives

My evaluation is summarized as responses to the stated objectives for the Results presented in this report. Those are re-stated below (from page v of the Executive Summary) and each is followed by my terse assessment.

1. ***Explore ways to summarize complex analyses and results into graphs that are easy to understand, interpret, and explain to decision-makers.***

The report assumes that the reader knows a lot about the history of this project. It also assumes that the reader has a decent understanding of statistical probabilities. I know very few real decision makers (including those who sit on the panels or boards of regulatory agencies) who meet those expectations. Those I know are politically astute and understand how to deal with "the odds", but I challenge you to fill a table with powerful people who will make a scientifically sound decision when, for example (page 53), they are told:

> "*The recovery standard is met when the geometric mean of projected escapement for the sixth best Snake River index stock over the last 8 years of a 48-year period exceeds the pre-determined recovery escapement level an average of 50% of the time...*"

I think that their eyes will glaze over, those still awake will shuffle papers in displacement activity, and the decision making process will derive from some combination of factors other than those being conveyed by well-meaning scientists. There has to be a better way. I strongly encourage you to distill the process to the point where there are only a handful of choices. Two or three is probably best. For all its machinations in alternatives and combinations of them (all 5,148!), the primary and practical choices actually distill to the three you identify as A1, A2 or A3. Offer those options up front.

*2. **Provide preliminary insights into the relative performance of alternative actions.***

This report certainly accomplishes that objective. It doesn't identify key elements in the causal chain because it doesn't deal with all components of the whole system. Those outside the river component are treated as unknowns. On the other hand, management within the river habitats is the only practical alternative available. If that doesn't prove successful, then the problem derives from the ocean habitat and that is an even more formidable challenge.

As a general observation, the results of these modeling analyses suggest only modest promise for the recovery of these salmon stocks. Strong encouragement (high probability) comes from only the most stringent measures applied to the assumptions and functions built into FLUSH. The sensitivity analyses suggest that FLUSH offers strong contrast with the forecasts of CRiSP when employed to test the suite of management actions expressed in A3. In my view, this identifies the most parsimonious test of the alternatives. Implementing the components of A3 will give the greatest prospect for resolving the unknowns. If the predictions of FLUSH are supported as a result of A3 management actions, then you've found the path to recovery. If they are not, then dispose of that as a simple approach and move toward evaluating the remaining hypotheses. I advocate an active, aggressive approach through experimental manipulation as the most responsible and expeditious way to resolve these controversies. A carefully designed management program can quickly reduce the dimensions of the problem.

Walters (Conservation Ecology, in press) has recently reviewed the many attempts at adaptive management and concludes that most have failed because their implementation was a compromise imposed on rigorous experimental design. That is a stern warning for this program. If an experimental management effort is to be developed, it must include commitment to strong, sustained manipulation. Replication and/or reference or control conditions are essential. Given that N = 1 for this problem, two conditions must be met: strong treatment effects plus good pre- and post-treatment data.

*3. **Identify key uncertainties that affect the results.***

The uncertainties are extensively considered within the constraints imposed by bounding the system as that between the nursery habitat and that above Bonneville. However, as made apparent in the accompanying document ("Recent Analysis of D Values"), we do not see a consensus on assumptions or analyses that precede the quantitative evaluation of uncertainty. In other words, we remain ill-informed and, therefore, quite uncertain about what the models can tell us let alone what we might learn from a formal analysis of uncertainty. This is confusing and discouraging in terms of moving toward management actions, but is encouraging evidence of a rigorous scientific exchange. Again, resolution to this disagreement will be best derived from a management experiment.

4. ***Test the sensitivity of decisions to the weights placed on key uncertainties, so as to focus the assessment of existing evidence, and the acquisition of additional evidence.***

Yes, the analysis based on weighting is extensive. Although, the core assumptions appear to be open to question (see comment #3). How the weightings are derived continues to be a subjective process. The qualitative approach offered in Table 5.7-1 may turn out to be as good as any. That lets decision-making develop as something akin to a Delphi process. That's good in that it builds consensus. That's questionable in that it will tend to disfavor strong alternative views--the kind that produce rigorous tests of alternatives. Each of the Normative River characteristics will have some fans and some opponents. The final subset of those will tend to be the ones considered least objectionable and/or most strongly proclaimed by a few individuals. That serves political purposes well, but may not produce the best scientific result. Again, designing and conducting a management experiment would be my first choice for resolving uncertainty.

5. ***Summarize results for some other important performance measures.***

This is a progress report. There's more to do here. I would be particularly interested in seeing some effort directed toward discounting much of the early evidence used in the retrospective analyses. A Bayesian approach can do that if you let it put more credence in recent evidence and diminish the role of that of the distant past.

There's so much reason for disagreement about what the data have to say up to about the mid-80's that a "best model" contest might be more profitably staged for the period when people have greater confidence in the monitoring effort. There are various effects of history embedded in the more recent data, but the PIT tag results are least arguable and, therefore, a more solid basis for evaluating the survivorship rates in recent years. In other words, by creating the expectation that a model will be evaluated in its ability to do retrospective analyses, we assume that the evidence from the past is of equivalent quality for each set of observations. That is highly unlikely. In addition, the multiple causality problems of changing collection and transport practices, the advent of a predator population around dams and the predator control program, changes in hatchery practices (i.e., BKD effects), the regime shift(s), etc., all provide further confounding effects. While it's true that we can't forecast future weather or climate or the applicability of our mechanistic understanding in that new context, we can minimize the constraint of bad or biased evidence in constructing that "prospective" forecast. Consider the arguments in "Recent Analysis of D Values" and how they might be reduced if we simply focused on data since 1978 or 1985. The model results seem more concordant after that or, when not in agreement, can be distinguished by direct challenges to the alternatives. I don't know how to make that recommendation more specific. That would have to come from an agreement among the empiricists and the modeling groups.

Other Issues

1. The detailed treatments expected in Appendix A suffer from missing parts. That's distracting but important in only one major case. I believe that the attention given to predation effects is sorely lacking. We know that predation is a major cause of mortality in juvenile fishes. We know that squawfish are the major piscivore. The squawfish removal program is a bona fide experimental treatment. Its effects can and should be dealt with in greater detail. Given all the other unknowns about screens, transport, stress effects, etc., it would seem advantageous to focus on one change that has been and can be readily assessed as a treatment effect because it has been recently applied and it could be removed. Squawfish aggregate above and below dams, making their effect more concentrated in space and time. The program has an assessment in place. Add estimates for the other piscivores (smallmouth bass and walleye) and that allows a reduction in one of the unknowns.

Be aware, however, that the majority of the squawfish taken by bounty fishers derive from below Bonneville. That has two implications. First, that section of the river has been and may continue to be an area of major losses to predators. Second, it is not among the reaches that can be effectively assessed using current approaches such as those employed at the dams.

There is a contradiction apparent in the fact that we seem to have very limited success in restoring salmon stocks, yet salmon appear to be quickly and successfully naturalized to wholly novel conditions. For example, chinook have developed substantial and self-sustaining populations in places that are very different from the habitat requisites usually ascribed to this species. Chinook have naturalized--and their populations quickly expanded to the limits of nursery habitat-- in many streams of the Great Lakes region, including Lake Superior which is among the most oligotrophic lake environments. Chinook have succeeded in establishing self-sustaining populations in many areas of New Zealand where their natal streams have virtually no estuarine habitat. How do they manage to accomplish these successes in places that clearly offer only sub-optimal habitats? My guess is that they do so because there are few natural predators in place to prevent that. It's only a guess, but there don't seem to be more reasonable explanations. All the hand-wringing about unknown physiological stressors doesn't fit as a tenable explanation for death of smolts. That leads me to conclude that a stronger focus on predation effects would improve the PATH process.

In addition, I wonder about the role of shad as a competitor and as a prey that subsidizes predator populations to levels that increase their effect on smolts. These salmon now migrate through a resident fish community very different from that of their past. I see little evidence of that in the considerations of these models. In the Great Lakes region we have learned that assertions about miscellaneous stressors pale by comparison to the obvious causes of sustained change--habitat destruction, invasion by exotics and predation effects. Those are the parsimonious explanations.

2. Given all the speculation and argument about regime shifts and climate effects, consider the learning opportunity offered by this year's El Nino. If variability in ocean conditions are, in fact, a major contributor to mortality then this year will have offered a strong contrast and, therefore, a specific basis for testing extra mortality hypotheses. The most recent regime shift appears to have begun well before this El Nino and its effects could be separated and independently evaluated.

3. I am concerned about semantics as they play out in the decision-making process. We advocate and use "uncertainty" as a code word among fishery scientists. That's not a problem in this context, but can become a problem at the levels where regulatory decisions get made. In that context, assertions of uncertainty are too often perceived as admissions of ignorance, indecision and disagreement. The scientists don't know what they're doing and they can't agree. This encourages decision makers to avoid decisions other than the obvious choice-- make choices on political criteria and defer hard choices until the scientists get things sorted out. Forthright decisions are essential to the pursuit of management manipulations required to test the alternative hypotheses. This report summarizes an effort important in attempting to critically evaluate complicated interactions, but the result is that we are awash in alternatives. I strongly encourage the PATH team to work toward couching its recommendations in ways that reflect two kinds of choices: high risk-high payoff vs. low risk-low payoff. We cannot make more than recommendations, but we can present those in ways that give a clear choice about the political will required to implement changes and, thereby, improve the chances for salmon restoration.

4. Lastly, I would remind the scientists that our Endangered Species Act is not a Law of Nature. It is an act of Congress. That can change with the next election or it can change as the public realizes the trade-offs required to sustain the Snake River chinook stocks or it can change as the courts interpet Treaty Rights challenges. As you approach the hatchery questions in the next iteration, be prepared to include the alternative favoring hatchery production as the only way to sustain salmon runs that meet public expectations and the requirements of Treaty obligations. That is an undesirable option in most circles, but it may be the only one. Considering its costs and benefits now might help emphasize the merits of the restoration alternative. We've had to accept that concession in the Great Lakes. Pacific salmon are no longer deemed exotics here. They're classed as naturalized and resident species. Restoring the native fish communities is no longer a singular goal. We must learn to manage around a mix of species interactions and relative abundances very different from those of the past. Public pressure and political pragmatism caused that. Be prepared for similar developments in your neighborhood.

Dr. Saul Saila

Reviewer: S.B. Saila
Title of Paper: Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook
Author: Marmorek and Peters [eds.]

Comments:

a) Scientific soundness of the analytical methods used

In general, I believe that the retrospective and prospective analyses, as well as the decision analysis framework are scientifically sound. However, I believe that there may be some areas which deserve further careful consideration.

1) For example, it is thought that the role of uncertainty in the identification of models for spring/summer chinook salmon, and in the application of these models to prediction of management alternatives may be underestimated.

   Four problem areas are mentioned below. They are:

   1) Uncertainty about model structure;
   2) Uncertainty in the estimated model parameter values;
   3) The propagation of model prediction errors, and
   4) The design of experiments in order to reduce the critical uncertainties associated with the models.

   Although I believe progress has been made in all of the above, it is my opinion that there is still much uncertainty about the variables characterizing the dynamic behavior of the models (Point I above). However, my major concern still is Point 2 above-the uncertainty about the values of parameters (coefficients) appearing in the identified structure of the dynamic models for the behavior of the fishery.

2) Although my knowledge of decision analysis is very limited, I would like to offer some comments and possible suggestions. From my respective, the decision-analysis process seems to consist of several steps which are addressed in an iterative fashion. The process is believed to include the following:

   1) Problem definition
   2) Identification of objectives and alternatives
   3) Descriptive and modeling of the problem, including model structure, uncertainty, and preferences
   4) Choice of alternative(s)
   5) Sensitivity. analysis
   6) Implementation of chosen alternative(s).

Some questions have occurred to me regarding the PATH implementation of the decision-analysis process.

a) Have the actual managers and decision makers been sufficiently involved in problem definition and identification of objectives and alternatives? It is my opinion that careful consideration of all aspects of the problem by a diverse group can lead to choosing alternatives that are not obvious to others.

b) It is my understanding (admittedly based on limited knowledge) that decision trees can be difficult to interpret (messy) as problem complexity rises. I think this one has arrived! In presenting results to managers, it seems very important that managers and decision makers receive a compact and completely understandable representation of the decision analysis. I believe that influence diagrams may be a superior approach to provide understanding to people regardless of their mathematical competence. For further information on this suggestion, see for example, Schachfer, R. 1986. Evaluating influence diagrams. Operations Research 34:871-882.

3) A final question arises about ensuring that all elements of the decision process are clearly defined. It seems to me that there still are disparate ideas regarding some aspects of the decision-analysis process. A possible resolution to this problem may be the so-called clarity tool (Howard, R.A. 1988. Decision Analysis and Promise. Management Science 34:679-695). This approach has been used to refine the conceptualization of events and variables associated with decisions so that the problem is defined well enough so that everyone can agree on definitions of the basic decision elements.

b) Reasonableness of the hypotheses incorporated into the analyses, the evidence on which they are based, and the modeling approach used to represent them

The three alternative hydrosystem actions (Al, A2, and A3) may have been too narrowly defined because I believe the actual decision makers and managers were not sufficiently involved with the initial problem identification and the identification of objectives and alternatives. In my opinion, the evidence on which the hypotheses and the modeling approach used were based, primarily on scientific evidence and did not necessarily include adequate inputs from the ultimate decision makers and managers. I believe this omission may have inhibited the discovery of alternatives that were not obvious to a more restricted group of biologist and modelers. The modeling approach has primarily focused on decomposing problems to understand their structures and to measure uncertainty. Decision trees have been used to create a model of the decision problem. Sensitivity analysis has been performed to answer “what if” questions.

The primary advantage of an appropriate model from a decision making perspective is believed to be that the mathematical representation of the problem substantially aids the decision maker to identify a preferred" alternative. I believe that this advantage has not yet been adequately achieved to date with the modeling process.

c) Validity of inferences and conclusions reached

The validity of inferences and conclusions reached are affected by several factors, and it is believed these factors may have adverse effects on them. For example, the inherent uncertainties in the spring/summer chinook population dynamics and system model parameters, reduces the validity of inferences and conclusions. It also seems evident that different perspectives regarding fish passage lead to different conclusions. In my opinion, this source of difficulty is particularly pertinent when more than one system submodel (CRISP vs. FLUSH) is used in developing inferences and conclusions. However, I believe that a genuine effort has been made to accommodate diverse perspectives. This is good, but I believe that formal efforts at resolving these issues will facilitate the ultimate decision making process.

d) Suggestions for improvements and extensions to the analytical approaches used

This category is utilized to provide many of the observations and suggestions which has come to mind for this reviewer. They are listed as suggestions which are hopefully constructive and refer to the entire decision-analysis process. Some of these are also included in other sections of this review. In this section, reference is made to the section of the report for which the comment or suggestion is considered to be relevant.

- page ii, Executive Summary, third paragraph-I believe there are objective methods for making consensus on the likelihood of various alternatives by means of so-called conflict resolution algorithms.
- Page ii, Executive Summary, Decision Options, end of first paragraph-I believe that all of the actual managers and decision makers related to spring/summary chinook should be more actively involved in the evaluation of the biological decision analysis and modeling tools. It seems to me that it is highly desirable to involve this group in order to broaden the available options.
- General - I believe that the effects of temporal changes in habitat quality as a result of dam construction should receive more careful consideration.
- Page ii, Decision Options, last paragraph-It seems to me that it is necessary to initiate the careful planning and design of experiments very soon in order to help resolve uncertainties which are insufficiently resolved under current conditions.

- Page iii, Uncertainties in the Response of Populations to Management Actions, first paragraph-I believe that some resolution of the relative importance of factors affecting stock survival can be achieved by means of the so-called clarity test. See, for example, Howard, R.A. 1988. Decision Analysis: Practice and Promise. Management Science 34:679695. It seems that much of the difficulty in decision making areas when different people (organizations) have different ideas regarding some aspect of the decision. The solution is to refine the conceptualizations of events and variables associated with the decision enough so that it can be made. I believe the clarity test provides a simple and understandable answer.
- Page iv, Item No. 12-Can the information derived from a recent drawdown and flushing of the Colorado River help in better defining the temporal scale of drawdown effects?
- Page 25, 4.2.1, Passage Models-It seems evident that the CRISP model involves substantially more parameters than the FLUSH model. I believe that the mechanistic approach involving several parameters may contribute substantially to the perceived differences between model outputs. In this case. I believe that parsimony in the number of uncertain parameters leads to a more realistic representation. Page 41, beginning of second paragraph-I have some questions regarding the reasonableness of defining habitat effects in terms of the Ricker α parameter.
- Page 42, 4.3.6, Hatcheries-It is my belief that the relatively rich literature on Atlantic salmon might help resolve the effects of hatchery releases on wild stock spawner-recruit survival.
- Page 45, 4. The Mainstem River ...-The fact that quantity and quality of mainstem rivers habitat is not considered is believed to be a significant omission. Limiting future conditions in the mainstem to fish travel time and migrational survival is considered to be unrealistic. Page 49, 5.2, Ways to summarize results: last paragraph-In my opinion, another possible approach to summarizing results of decision analysis includes the following. Screening alternatives on the basis of dominance is an important decision-analysis tool. To this end, I suggest that cumulative risk profiles for alternatives be given serious consideration. See, for example, Bunn, D. 1984. Applied Decision Analysis. New York, McGraw Hall, for details regarding both deterministic and stochastic cumulative risk profiles and their applications.
- Page 51, 5.2.3, Expected ability of actions to meet some criterion, first paragraph - In my opinion, the preferred way to weight each aggregate hypothesis by a probability value is by the use of decision analysis computer programs. Sources for suitable computer programs include the following. Supertree is a full-feature decision analysis system developed by SDG Systems, Menlo Park, California. The address for Supertree is:

SDG Systems, Inc.
3000 Sand Hill Road
Menlo Park, CA 94025

Other decision analysis software include:

Arborist	Texas Instruments, Inc. PO Box 1444 Houston, TX 77251
In Dia	Decision Focus, Inc. 4984 El Camino Real - Suite 200 Los Altos, CA 94022

e) Opportunities for using an adaptive management approach to resolve some of the remaining uncertainties

In general, I believe there are many opportunities for reducing some of the uncertainties by an adaptive management approach. My perception of an adaptive management approach involves systematically varying management options while carefully monitoring biological, economic, and social consequences of the actions, in an effort to reduce uncertainty and to apply the new information to quantitative models.

Hilborn and Walters (1992, Quantitative Fisheries Stock Assessment, page 491) outline six steps to adaptive management which included:

1) Alternative hypothesis identification;
2) Estimating the expected value of perfect information to determine if further steps are acquired;
3) Model development for future learning;
4) Identification of adaptive policy options;
5) Performance criteria development; and
6) Formal option comparison by decision-analysis techniques.

It is my opinion that not enough attention has been given to inputs from actual decision makers and managers regarding Step I above. I also believe no explicit steps have been taken to estimate the expected value of perfect information, which provides an upper bound for the expected value of information in general. I believe that a decision tree can be redrawn so that the uncertainty nodes for which perfect information is available come before the decision node.

f) Relative priorities for future work on these analyses

It is evident that my comments and suggestions have centered on the decision-analysis aspects of this PATH report, and these are scattered through my report.

I think that because influence diagrams provide a relatively simple and intuitively comprehensible graphical representation of a decision problem, they could be helpful in structuring decision to non-specialists.

General Comments

It is my belief that although simulation models, such as those applied in the PATH analysis, can represent systems of realistic complexity, they are limited by factors which arise from the way in which they are built. The usefulness of these system models is ultimately limited not by computer technology but by our knowledge of the system dynamics and the effects of management practices upon them. The acquisition of such knowledge is a very slow and difficult process. It is suggested that alternative paradigms for modeling biotic systems based on- development in the field of artificial intelligence should also receive consideration in future PATH model development.

Dr. Carl Walters

Reviewer: Carl Walters
Title of Paper: Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook
Author: Marmorek and Peters [eds.]

General comments:

(1) I am truly impressed with the way you have pulled together hypotheses and uncertainties into an overall framework for decision analysis. The approach of sampling from a large set of alternative hypotheses to provide probability distributions for outcomes is an important general contribution beyond its application in the Columbia, and it makes very clear that there are a few really important "leading uncertainties" that cannot be resolved by small-scale research.

(2) There is an "imbalance" that my experience suggests will come back to haunt you in the near future. You have done the usual biologist's thing of articulating a complex set of biological response hypotheses, but you have articulated only a few policy alternatives. On reflection, the few adaptive management success stories have involved the opposite: relatively few response hypotheses, but a very rich set of policy alternatives. Perhaps your experimental management workshops will begin to expose that set. As one of the fathers of systems thinking once said: "I don't need systems analysis, I need imaginative synthesis".

(3) Your basic result about low probabilities of recovery is not surprising, but it is certainly discouraging especially in relation to the expensive dam removal options. I worry that even this assessment may be systematically too optimistic, for two technical reasons:

(a) if I read the simulation methods correctly, you sample from historical flow/survival sequences. If this sampling procedure failed to preserve autocorrelation structure in those sequences (fewer runs of extreme conditions in simulation samples) you could greatly underestimate resulting variance in population sizes over time; check this and explain more precisely in the report text.

(b) Your simulation trials should each represent a sample from the joint posterior for parameter values, which is different from sampling from the hypotheses with equal odds. For example, when you have chosen a river survival model (CRiSP vs FLUSH), the population parameters for simulation should be those consistent with that model (e.g., estuary/marine survival rate consistent with the S/R data given the survival model chosen). It is not clear in the writeup whether you did restrict the parameter sample this way; doing so would tend to reduce the variance in outcomes a bit, but not enough to counter effects of incorrect correlation structure.

(4) I was pleased to see use of "history reference" policy comparisons (what would have happened under different policy given historical patterns of variation); it is not clear how far you took this approach in the simulations--see comment above about importance of retaining correlation structure in all "random" time effects.

(5) It detracts from the value of your analysis to use the myopic, bureaucratic Jeopardy standards as measures for comparison of alternatives. Those standards may be convenient, but utterly fail to recognise the history of stock declines prior to 1950 (restoration to an already bad situation 1950-70 is hardly what the public is likely expecting). Further, they lead you to focus mainly on the stocks that were still relatively healthy as of 1970 before the last dams went in, i.e. your "weaker" or "weakest" index stocks are already a grossly non-representative subsample of the Columbia's long term potential. Why not produce some more comprehensive measures for restoration of biodiversity and system-scale productivity, by including simulations for a broader selection of stock "types" including some that were already too depressed to provide good S-R data for your baseline parameter estimation? The bottom line here is that in further refining/developing objectives for decision analysis, you need to get away from the presumption that the main audience for your results is the collection of NMFS and other agency people. Like them, your first responsibility is to the public, whether they remember this or not.